



UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Bachelor's Thesis in Financial Economics
15 ECTS

**Evaluating the Efficiency of the
Swedish Stock Market:**
a Markovian Approach

Authors:
Emil Grimsved
John Pavia

Supervisor:
Mattias Sundén

Spring 2015
Department of Economics

Evaluating the Efficiency of the Swedish Stock Market: a Markovian Approach

Emil Grimsved and John Pavia*

Abstract

This paper evaluates weak form efficiency of the Swedish stock market, by testing whether or not the index OMXSPI follows a random walk. The returns of the index are mapped onto one of two states, and the resulting data set is treated as a higher-order Markov chain for the purpose of analysis. The Bayesian information criterion is used to determine the optimal order of the chain and the null hypothesis that the chain is of order zero is tested against the alternative that the chain is of the established optimal order. We find that random walk behaviour cannot be rejected for the period January 2000 to April 2015.

JEL classification: C60, G14.

Keywords: efficient market hypothesis, EMH, random walk hypothesis, RWH, Markov chains, OMXSPI, time dependence, time homogeneity, weak form efficiency, Bayesian information criterion.

HOW DO YOU BEAT THE MARKET? For obvious reasons this is perhaps the single most important question in portfolio management. The question suggests that outperforming the market is possible, which would mean that investors consistently can earn returns that are higher than the expected market return.

Advocates of the efficient market hypothesis (EMH) disagree. In fact, EMH directly implies that it is impossible to develop a trading strategy that consistently beats the

*School of Business, Economics and Law, University of Gothenburg. We would like to thank Mattias Sundén for his comments, feedback and support, which have been of great value to us. Throughout this report MATLAB has been used. The code and the corresponding data can be sent upon request through mailing to john.emil.thesis@gmail.com.

market over time (Malkiel, 2005). This does not mean that investors cannot beat the market, but that if they do so, it is not due to fact that they have a superior trading strategy; they simply owe their success to chance.

The efficient market hypothesis says that capital markets are efficient, which means that all available information that is relevant to the pricing of an asset is incorporated in the price of the same asset (Fama, 1970). The concept of market efficiency is closely related to the idea that the movements of stock prices are indistinguishable from those of a random walk. This idea is known as the random walk hypothesis (RWH). The two hypotheses are related in the sense that if stock prices indeed follow a random walk, future stock prices cannot be predicted, and hence no trading strategy that consistently beats the market can be developed (Malkiel, 2005).

The opinions on the degree of efficiency in stock markets differ among financial economists, and many believe that there may be some degree of predictability in stock market returns (cf. Fama & French 1988; Malkiel, 2003; Schiller, 2014). Even so, the question of whether or not potential patterns in returns can be exploited profitably, as well as market efficiency as a general concept, are still highly debated topics within the field of financial economics.

Market efficiency is not just of interest as a theoretical concept; it also has implications for the actions of market participants. In an efficient market all available information about an asset is reflected in its price, and thus market efficiency is of obvious interest since it ensures that prices give accurate signals for investment decisions (Fama, 1970).

This paper aims to evaluate the efficiency of the Swedish stock market, by testing whether or not the price of the index OMXSPI follows a random walk. This means that our main interest is to test the Swedish stock market for what Fama (1970) refers to as weak form efficiency, which in turn means that the question of interest is whether or not information about historical prices are incorporated in the current price of the index. Many such test, for various markets, have been performed (cf. Fama, 1970, 1991, 2014), including tests for the Swedish stock market (cf. Frennberg & Hansson, 1993; Shaker, 2013). The methodology has differed between the tests and for the Swedish stock market mostly variations of autoregressions have been the models of choice. This paper develops a Markovian model inspired by the one used by Fielitz and Bhargava (1973) as well as the one used by McQueen and Thorley (1991). The model is based on the fact that independent returns is a sufficient condition for a random walk in prices, and hence random walk behaviour can be tested by assessing the dependence structure of returns. A data set consisting of indicator variables representing high and low returns respectively is constructed and tested for dependence structures by estimating transition probabilities under the assumption that the data set represents a Markov chain of a given order.

The Markovian model has several advantages over an autoregressive one. It is non-parametric, and hence no assumptions about the distribution from which the data is sampled have to be made. The model also allows for non-linear dependences (McQueen & Thorley, 1991), as the transition probabilities are allowed to vary depending on previously

realised returns. Additionally, since the returns are mapped onto states, the model is insensitive to outliers, and therefore the whole sample can be used for the purpose of analysis. These advantages come at a price, as the Markovian model requires other strong assumptions; the chain representing the returns must be aperiodic, irreducible and time homogeneous. The first two of these assumptions will be validated in the estimation procedure, while the last one will be tested explicitly.

This paper offers an extension of previous models used to test random walk behaviour in stock market prices using Markov chains (cf. Fielitz & Bhargava, 1973; McQueen & Thorley, 1991). Instead of fixing an order of the chain, and thereby limiting the analysis to a certain dependence structure, an optimal order is derived. Further, it is shown that pairwise tests cannot be used to reliably establish the optimal order of the chain, and therefore an information criterion, namely Bayesian information criterion (BIC), is used instead. In addition, the paper contributes with a discussion on the highest testable order of the Markov chain, and outlines the test for time homogeneity in detail. Finally, in the test for time homogeneity a correction of the degrees of freedom of the test statistic is presented, as previous papers have either been unclear or fallacious in this particular matter (cf. Fielitz & Bhargava, 1973; Tan & Yilmaz, 2002).

In terms of delimitations, the state space on which the Markov chain is defined only consists of two states. In addition, the order of the chain is not allowed to vary over the time period. This is mainly due to time constraints, but also due to the fact that the focus is devoted to extending the fixed-order Markovian model to improve the reliability of the test. Further, only the efficiency of the Swedish stock market, represented by the index OMXSPI, is evaluated. Data of different frequencies is, however, considered. Daily, weekly as well as monthly price data are analysed as there may be different dependence structures in the different data sets.

To summarise, the questions this paper attempts to answer are:

- Do the prices of the Swedish stock market during the period January 2000 to April 2015 follow a random walk? Equivalently, can the returns during the period be modelled by a zero-order Markov chain?
- Is the assumption of time homogeneity of the Markov chain reasonable?
- Based on the results from the test for random walk behaviour, can the Swedish stock market be considered to be weak form efficient?

The remainder of this paper is structured as follows: section I reviews previous research on random walks and efficient markets. Section II is devoted to the efficient market hypothesis and the random walk hypothesis. In section III, which is rather technical, the methodology for testing whether or not the index OMXSPI follows a random walk is developed. The data sets of interest are described in section IV. Finally, the results and the subsequent discussion are given in sections V and VI, respectively.

I. A Review of Past Results

Over the last fifty years, many tests for random walks in stock prices and market efficiency have been published. This section presents an overview of what has been done within the field of market efficiency and on random walks in asset prices. The overview is limited to studies that either have used a Markovian approach or where the market of interest has been the Swedish stock market.

Most studies that have tested for random walks in stock prices using a Markovian model have employed it on the US stock markets. For example, Niederhoffer and Osborne (1966) rejected random walk behaviour of a set of stocks traded at the New York Stock Exchange (NYSE) when considering intraday returns modelled by a second-order Markov chain. These results were confirmed by Fielitz and Bhargava (1973), who used a first-order Markov chain to model returns of a set of stocks. Fielitz and Bhargava included three states, which allowed them to model magnitudes. A random walk in stock price was rejected for the vast majority of the stocks. In a paper from 1975, Fielitz again used a Markov chain of order one to test for time dependence in returns of individual securities traded at the NYSE. For short time periods it was found that there existed a weak price memory, which means that returns could be predicted for short time periods.

McQueen and Thorley (1991) used a second-order Markov chain to test the random walk hypothesis using annual returns from the NYSE. They found that the real prices of the NYSE showed significant deviations from random walk behaviour. These results confirmed findings by Lo and MacKinlay (1988). McQueen and Thorley did not test the assumption about time homogeneity of the Markov chain.

Tan and Yilmaz (2002) presented criticism against the methods that McQueen and Thorley, and Fielitz and Bhargava used. The criticism was based on McQueen and Thorley's failure to test if the assumptions of the model, in particular the assumption about time homogeneity, held. In addition, Tan and Yilmaz criticised Fielitz and Bhargava for performing tests that required the Markov chain to be time homogeneous, even though they had rejected the same assumption.

The research on market efficiency and random walks in Swedish stock prices is limited. No study has used a Markovian model to examine random walk behaviour of the Swedish stock market, but other methods such as autoregressions, variance ratio and serial correlation tests have been used (cf. Jennergren & Korsvold, 1974; Frennberg & Hansson, 1993; Shaker, 2013).

The Swedish and Norwegian stock markets were tested for random walk behaviour by Jennergren and Korsvold (1974). They considered 45 stocks, and rejected a random walk behaviour for a majority of those. Frennberg and Hansson (1993) tested and rejected random walk behaviour of the Swedish stock market for the time period 1919 to 1990. They confirmed findings from the US stock markets (cf. Lo & MacKinlay, 1988; Poterba & Summers, 1988) where returns over long periods exhibited mean reversion, while short

horizon returns showed positive autocorrelation. By using Swedish stock market data from the time period 1986 to 2004, Metghalchi, Chang and Marcucci (2008) tested three different trading rules based on a moving average. They found that these trading rules could outperform a simple buy and hold strategy even if transaction costs were included. Shaker (2013) examined the random walk behaviour of the Swedish stock market using daily closing prices of the index OMXS30 during the time period 2003 to 2013. He rejected both weak form market efficiency and random walk behaviour of the Swedish stock market using variance ratio and serial correlation tests.

II. Efficient Markets and the Random Walk Hypothesis

This section introduces the efficient market hypothesis and the random walk hypothesis. The section starts with a presentation of the efficient market hypothesis and discusses how market efficiency can be evaluated. An introduction to the random walk hypothesis follows. The section concludes with a discussion about the relationship between random walks in stock prices and the efficient market hypothesis.

A. *The Efficient Market Hypothesis*

Market efficiency has been a highly debated subject in economic theory ever since Eugene Fama presented his doctoral dissertation in the 1960s. A market is said to be efficient if all information that is available and relevant to the pricing of an asset is incorporated in the price of the same asset (Fama, 1970, 1991). The efficient market hypothesis (EMH) then simply says that stock markets are efficient in the described sense (Fama, 1970). The term efficiency itself refers to the idea that a market with the described property gives "accurate signals for resource allocation" (Fama, 1970, pp. 1), thus making capital markets efficient.

A necessary condition for this strong version of EMH is that there are no transaction costs, nor any expenses related to the acquiring of relevant information. Weaker versions of the hypothesis, which have the benefit of being more economically reasonable, have been suggested. Jensen (1978) introduced a version where a market is efficient if the marginal benefit of acting on information is no higher than the marginal cost of the same action. In other words, by this definition, prices only need to reflect information on which it would otherwise have been profitable to act.

Testing EMH is not possible unless which information set is used is specified (Fama, 1970). To make the hypothesis testable Fama introduced three types of tests corresponding to three subsets of information (Fama, 1970, 1991): weak form tests, where the information set consists of historical security prices and other market observable variables; semi-strong form tests, where the information set also includes other publicly available information; and strong form tests, where private information is included as well.

In 1991, Fama changed these categories into ones that says more about what is actually tested for. Weak, semi-strong and strong form tests were now introduced as tests for return predictability, event studies and tests for private information, respectively.

Tests of market efficiency relates observed prices to equilibrium prices in the sense that under EMH the observed price should exhibit the properties of the equilibrium price (Fama, 2014). The efficient market hypothesis thus has to be tested jointly with an asset pricing model, which is used to model equilibrium returns or prices. If the specified equilibrium asset pricing model does not hold, efficiency may be rejected because of an inadequate specification of the returns even though relevant information may be incorporated in prices. In general there is no way of determining if market inefficiency, the pricing model or some combination of the two is the reason for the rejection (Fama, 2014). This difficulty, known as the joint hypothesis problem, makes the choosing of a reasonable pricing model a crucial part of testing EMH.

B. The Random Walk Hypothesis

The theory of random walks in stock prices dates back to 1900 when Louis Bachelier presented his dissertation *The Theory of Speculation*. Fama (1965, pp. 56) defines a market to be a random walk market if "successive price changes in individual securities are independent". If price changes are independent, and transaction costs are ignored, complicated trading strategies will not be more successful than a simple buy and hold strategy, since the price development of securities cannot be predicted.

The notion that price development is unpredictable is consistent with the random walk hypothesis (RWH), which says that the movement of stock prices cannot be distinguished from a those of a random walk (Fama, 1965; Malkiel, 2005). This is the same as to say that the development of a partial sum of a sequence of independent random numbers is equally unpredictable as the future path of the asset prices. According to Fama (1965), the random walk hypothesis is not an exact description of real asset price behaviour. Even so, the dependence structure may be weak enough to consider RWH to be a reasonable approximate description of the movements of stock prices (Fama, 1965).

C. Random Walks and Efficient Markets

If the movements of stock prices are indistinguishable from those of a random walk investors cannot possibly predict returns and hence the efficient market hypothesis is associated with the idea that stock prices follow a random walk (Malkiel, 2003). It would be misleading to talk about any strict logical implications. The market could follow a walk because investors choose assets at random. While this is not likely, it illustrates that a random walk in prices is not a sufficient condition for market efficiency. Conversely, in the context of this paper, as returns are divided into states one may find that one can predict the direction of stock price movements, but not the magnitude of a rise or a fall in price.

Therefore, a test of RWH may lead to a situation where something can be said about the behaviour of the stock market, but where it is still impossible to beat the market consistently.

How the efficient market hypothesis is related to the random walk hypothesis has been a highly debated topic in the field of finance (cf. Lo & MacKinlay, 2002; Malkiel, 2003). Lo and MacKinlay (2002) conclude that the relationship of RWH and EMH cannot be explained in terms of sufficiency and necessity. However, economic literature (cf. Fama, 1991, 2014), suggests that a random walk in stock prices is consistent with the efficient market hypothesis, and in many studies (cf. Fama & Blume, 1966; Jensen, 1978) EMH and random walks in stock prices are evaluated in the same context. In this paper random walk behaviour in stock prices will be considered to be an indication of market efficiency and, inversely, non-random walk behaviour will be seen as evidence, but not as proof, of market inefficiency. There will, however, be no deeper evaluation of the relationship between the two.

III. The Markovian Model

This section outlines the procedure to test for random walks in stock prices. First, the construction of the Markov chain¹ modelled is presented. Thereafter, the Bayesian information criterion (BIC) is used to determine the order of the constructed Markov chain modelled. The null hypothesis that the constructed chain is of order zero is then tested against the alternative that the chain is of the order established by BIC. This is called a test for time dependence. Finally, as the estimation of the transition probabilities requires that the Markov chain is time homogenous, a test for time homogeneity is given.

A. Returns and Benchmark Returns

Let P_t be the price of an asset at time t , $t = 0, \dots, T$. The return², denoted r_t , $t \geq 1$, during the period $t - 1$ to t is then calculated as:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}. \quad (1)$$

Hence, r_t is the percentage change from one time period to the next.

The two benchmarks that are used in this paper are the geometric return and the zero return. The geometric return is calculated as:

¹An introduction to the Markov theory that is used in this model can be found in appendix A.

²It is worth to mention that log-returns are commonly used in empirical financial economics. One crucial reason for this is that the logarithmic transformation make the data look more normally distributed. As the model presented in this paper does not require any normality assumption, the more direct approach of assessing returns, rather than log-returns, can be taken.

$$\hat{r} = \left(\prod_{i=1}^T (1 + r_i) \right)^{1/T} - 1, \quad (2)$$

where T is the number of observations, i.e. the sample size of returns, and r_i is calculated as in (1). When zero is used as the benchmark \hat{r} is equal to 0. In the construction of the state space in section III.B the expected return, $\mathbb{E}[r]$, is replaced by \hat{r} which represents the estimate of $\mathbb{E}[r]$.

B. Mapping Returns to States

To model returns by a Markov chain which is discrete in both time and space, the returns have to be divided into states. This is done by assigning a rule that maps the returns onto the states on which the Markov chain is defined. In this paper two states are considered. Let $\{r_t, t = 1, \dots, T\}$ be a time series of returns. The returns are classified as "low" and "high" depending on whether or not the return is above the expected return $\mathbb{E}[r]$. Let the state space, S , consist of the two states L and H which indicate low and high returns, respectively. The returns are then mapped into this state space as follows:

$$X_t = \begin{cases} L & \text{if } r_t < \mathbb{E}[r] \\ H & \text{if } r_t \geq \mathbb{E}[r]. \end{cases} \quad (3)$$

Since, $\mathbb{E}[r]$ in (3) is unobservable it is replaced by any of the two benchmarks denoted by \hat{r} .

It would be possible to consider more than two states and have each state represent an interval within which the realised returns lie. The main reason for not using more than two states in this thesis is the difficulty of finding an unambiguous way of constructing such a mapping. This is, to a certain extent, true for two states as well, but at least two states are needed for the chain to carry any information at all.

C. Estimation of Transition Probabilities

The transition probabilities of a u :th order Markov chain are estimated under the assumption that the chain is time homogenous. Considering a time homogenous chain, the maximum likelihood estimates of the transition probabilities are given by³:

$$\hat{p}_{ij} = \frac{n_{ij}}{n_{i.}}, \quad \forall i \in S^u, \forall j \in S, \quad (4)$$

³A derivation of the maximum likelihood estimates of the transition probabilities can be found in appendix C.

which are obtained by maximising the likelihood function subject to the constraint: $\sum_j p_{ij} = 1, i \in S^u, j \in S$. The counts n_{ij} and n_i denote for the number of transitions from $i \in S^u$ to a specific $j \in S$ and the number of transitions from i to any state $j \in S$, respectively. The observed counts are displayed in a transition count matrix (TCM). The transition probabilities are displayed in a transition probability matrix (TPM), and the estimated transition probabilities are displayed in an estimated TPM. Note that the estimation procedure requires that each row in the TCM must sum to a positive value, since otherwise the denominator in (4) would be zero and the expression would not even be defined.

For a Markov chain of order two defined on the state space $S = \{H, L\}$, the TCM and the estimated TPM are displayed in figure 1 below:

Figure 1.

An illustration of the TCM and the TPM of a second-order Markov chain.

The figure illustrates the transition count matrix and the estimated transition probability matrix of a second-order Markov chain defined on a state space, S , consisting of two the states L and H . The entries in the TCM, n_{LLL}, \dots, n_{HHH} are the observed number of transitions for the second-order Markov chain followed that given path. The numbers \hat{p}_{ij} are the estimated transition probabilities for the associated sequences.

		TCM				TPM	
Previous states		Future state		Previous states		Future state	
		L	H			L	H
L	L	n_{LLL}	n_{LLH}	L	L	\hat{p}_{LLL}	$1 - \hat{p}_{LLL}$
L	H	n_{LHL}	n_{LHH}	L	H	\hat{p}_{LHL}	$1 - \hat{p}_{LHL}$
H	L	n_{HLL}	n_{HHL}	H	L	\hat{p}_{HLL}	$1 - \hat{p}_{HLL}$
H	H	n_{HHL}	n_{HHH}	H	H	\hat{p}_{HHL}	$1 - \hat{p}_{HHL}$

The TCM and TPM above can be generalised to a u :th order Markov chain defined on a state space with cardinality n_s in a straightforward manner.

D. Test for the Order of a Markov Chain

The aim of this section is to determine the order of the Markov chain modelled. Intuitively, multiple pairwise tests may seem appealing, and has previously been suggested by Tan and Yilmaz (2002). They presented the following procedure: the null hypothesis that the Markov chain is of order zero is tested against the alternative that the Markov chain is of order one. If the null hypothesis is rejected, the procedure is repeated, but this time order one is tested against order two. The pairwise tests continue until the null hypothesis that the Markov chain is of the lower order cannot be rejected, or until a specified highest

order, M , is reached. Whenever the test first fails to reject that the chain is of order $u \in \{0, 1, \dots, M-1\}$, when tested against the alternative that the chain is of order $u+1$, the chain is considered to be of order u .

However, it is possible, when testing a chain of order $u+1$, that the null hypothesis that the chain is of order $u-1$, cannot be rejected when tested against the alternative that the chain is of order u (see appendix C for further details). This shows that the procedure suggested by Tan and Yilmaz is not reliable.

Therefore, a more reasonable approach is to use an information criterion and in this paper the Bayesian information criterion (BIC) is used. The use of BIC when testing for the order of the chain can intuitively be motivated by the fact that it penalises for increasing the order of the chain if the additional information contained in the realisations of the added periods containing the additional information is insufficient. The main reason for choosing BIC⁴, over e.g. Akaike information criterion (AIC), is that the BIC gives both an optimal and consistent estimator of the order of the Markov chain. The use of BIC requires that a maximum allowed order, M , is specified in advance and a method for doing so is presented in section III.D.1. The procedure to estimate the order using BIC is given in section III.D.2.

D.1. Determining the Highest Possible Order

The method used to determine the order of the chain requires that a maximum order M is specified in advance. As it is possible, for any $u \in \mathbb{N}$, to construct a Markov chain that is of order u , but not of any order $v \in \mathbb{N}$ such that $v < u$,⁵ one cannot determine a highest order without considering the nature of the data set of interest. In the context of this paper, this would mean that one would have to present an argument for why it is economically unreasonable for a sequence of returns to be a Markov chain of an order higher than M .

There is, however, a technical limitation which must also be taken into consideration. For the test of the order of a chain to be valid it is required that each transition probability is strictly positive. This in turn implies that there must be at least one count in each entry in the corresponding TCM. For a given chain this means that once the order is high enough for the corresponding TCM to have an entry which equals zero, one must assume that the chain is not of this or any higher order.

In this paper the maximum order, M , will be set to the highest order that corresponds to a TCM whose entries are all non-zero. The motivation for choosing the maximum order M this way is simple. As it is a stronger assumption that a chain is of order $u \in \mathbb{N}$ than that the chain is of order $u+1$, and hence the larger M is, the weaker the assumption one has to make about the order of the chain becomes. By choosing M as above it gives the

⁴The Bayesian information Criterion is also known as Schwartz Bayesian criterion (SBC) since it was first derived by Schwartz (1978) to find the optimal dimension for the model used.

⁵In appendix C an example of such a construction is shown.

largest possible maximum order and hence also the weakest possible assumption about the order of the chain for each data set.

D.2. Test for the Order by Using an Information Criterion

This section describes a method, first presented by Anderson and Goodman in 1957, for deciding whether or not the TPM of a Markov chain of order $v < u$ is statistically different from the TPM of a chain of order u . The method is required to determine an order using the Bayesian information criterion (BIC). The order established using BIC is optimal, in the specific sense that, under the assumptions that the prior distribution is a non-informative Dirichlet distribution, it minimises the expected loss (Katz, 1981). The established order does not depend on either the prior distribution or the posterior distribution (Katz, 1981).

The BIC procedure requires that the state space, S , is finite and that the Markov chain is aperiodic and irreducible. Furthermore, as stated above, a maximum order, M , has to be specified. By determining M as above the assumptions of irreducibility and aperiodicity are fulfilled. Further, the state space $S = \{L, H\}$ is finite. Hence, the assumptions hold.

As for the testing procedure, which is based on the work of Anderson and Goodman (1957), consider a sequence of data which may be represented by a Markov chain. The objective is to test if the Markov chain is of order v against the alternative that the Markov chain is of order u . It can be assumed, without loss of generality, that $v < u$. In this setting there are three sequences to consider; $u = i_{n-u}, \dots, i_{n-1} \in S^u$ which carries information in the Markov chain of order u ; $v = i_{n-v}, \dots, i_{n-1} \in S^v$, which carries information in the Markov chain of order v ; and $d = i_{n-u}, \dots, i_{n-(v+1)} \in S^d = S^{u-v}$, which belongs to the set of sequences that separate the sequences in S^u from the ones in S^v , it follows that $u = dv$. The transition probabilities using this newly introduced notation for the chain of order u and the chain of order v are defined in equations (5) and (6), respectively:

$$p_{uj} = \mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_{n-v} = i_{n-v}, \dots, X_{n-u} = i_{n-u}) = p_{dvj}, \quad (5)$$

$$p_{vj} = \mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_{n-v} = i_{n-v}). \quad (6)$$

Let $n_{uj} = n_{dvj}$ be defined as the number of transitions following the sample path $i_{n-u} \dots i_{n-v} \dots i_{n-1}j$ for the Markov chain of order u . Analogously n_{vj} is defined as the number of transitions following the path $i_{n-v} \dots i_{n-1}j$ for the v :th order chain. Define $n_{u.} = n_{dv.}$ and $n_{v.}$ as the total number of transitions following the sample paths $i_{n-u} \dots i_{n-v} \dots i_{n-1}$ and $i_{n-v} \dots i_{n-1}$ for the Markov chains of order u and v , respectively.

Then the maximum likelihood estimates of the transition probabilities are calculated as in (4). However, using the notation introduced above, the transition probability given in (4) is now given by (7) and (8) for the chains of order u and v , respectively:

$$\hat{p}_{uj} = \frac{n_{uj}}{n_{u.}} = \frac{n_{dvj}}{n_{dv.}} = \hat{p}_{dvj}, \quad (7)$$

$$\hat{p}_{vj} = \frac{n_{vj}}{n_{v.}}. \quad (8)$$

The null hypothesis, H_0 , and the alternative hypothesis, H_1 , can be formulated⁶ as below:

H_0 : the chain is of the lower order v ,

H_1 : the chain is of the higher order u , but not of the lower order v .

The likelihood ratio statistic Λ_v for a given sequence, vj , is given below:

$$\Lambda_v = \prod_{v,j} \left(\frac{\hat{p}_{vj}}{\hat{p}_{dvj}} \right)^{n_{dvj}}. \quad (9)$$

There are n_s^{u-v} unique sequences in which the sample paths coincide. Therefore, the test statistic Λ becomes the product of the test statistics Λ_v . This is to say that:

$$\Lambda = \prod_d \Lambda_v = \prod_d \left(\prod_{v,j} \left(\frac{\hat{p}_{vj}}{\hat{p}_{dvj}} \right)^{n_{dvj}} \right) = \prod_{d,v,j} \left(\frac{\hat{p}_{vj}}{\hat{p}_{dvj}} \right)^{n_{dvj}}. \quad (10)$$

Taking the transform $-2 \log(\Lambda)$, the limiting result becomes:

$$-2 \log(\Lambda) = 2 \sum_{d,v,j} n_{dvj} \log \left(\frac{\hat{p}_{dvj}}{\hat{p}_{vj}} \right) \stackrel{a}{\sim} \chi_{df}^2, \quad df = (n_s^u - n_s^v)(n_s - 1). \quad (11)$$

The test statistic $-2 \log(\Lambda)$ under the null hypothesis is asymptotically Chi-squared distributed with $(n_s^u - n_s^v)(n_s - 1)$ degrees of freedom. Which is a generalisation of the test statistic derived in Anderson and Goodman (1957).

⁶The aim is to test whether or not the probability distribution of the u :th and v :th order Markov chains are the same. Mathematically, the null and alternative hypotheses can be stated as:

$$\begin{aligned} H_0 &: \forall u \in S^u, \forall j \in S; p_{uj} = p_{vj} \\ H_1 &: \exists u \in S^u, \exists j \in S; p_{uj} \neq p_{vj}. \end{aligned}$$

Under the assumptions that the Markov chain is aperiodic, irreducible and defined on a finite state space S , with an upper bound M of the order of the chain, the BIC estimator for the order of the chain is defined below.

DEFINITION 1: Let X be a Markov chain of order $u < M$. Let the likelihood ratio statistic, Λ , for testing order u versus order M be denoted by $\Lambda_{u,M}$, then the BIC estimator, \hat{u}_{BIC} , for the order of the Markov chain is such that:

$$f(\hat{u}_{BIC}) = \min_{0 \leq u < M} f(u), \quad (12)$$

where $f(u) = -2 \log(\Lambda_{u,M}) - (n_s^M - n_s^u)(n_s - 1) \log(T)$, T is the sample size and $(n_s^M - n_s^u)(n_s - 1)$ is the degrees of freedom for the likelihood ratio statistic $\Lambda_{u,M}$.

The likelihood ratio test statistic is at least as large when an order higher than u is tested against v , as it is when testing u against v . In the same sense as adding explanatory variables to a linear regression model never reduces the fit, adding periods that may carry information in the Markov chain never reduces the likelihood ratio statistic. Therefore, in the testing procedure the term $(n_s^M - n_s^u)(n_s - 1) \log(T)$ penalises for increasing the order, which can be compared to utilising the adjusted R -squared when additional explanatory variables are added to a multiple linear regression.

It should be noted that, when using BIC for estimating the order of a Markov chain, one tests the highest allowed order M against all lower orders $0, 1, \dots, M-1$ (Katz, 1981). If BIC gives the optimal order 0, then the BIC only says that this order best represents the data when penalising for the increased order. It does not determine whether or not there is any dependence structure in the returns. Thus, to be able to perform a significance test for time dependence, the optimal order, determined using BIC, must be at least 1. Therefore, we have chosen to test the maximum order M against the lower orders $u \in \{1, \dots, M-1\}$.

E. Testing if the Order of the Chain is Different from Zero

Assuming that an optimal order u has been established using BIC, u is the optimal order (or, rather, the optimal order different from zero) of the Markov chain, but BIC says nothing about whether or not this order results in a plausible model. If every order of the chain results in a bad model, BIC just gives us the least bad of these models. Therefore, a test must be performed to determine whether the order established using BIC results in a model that is significantly better than a Markov chain of order zero.

Under the assumption of time homogeneity, the null and alternative hypotheses can

be stated⁷ as below:

- H_0 : The chain of the optimal order is also a chain of order 0,
- H_1 : the chain of the optimal order is not a chain order 0.

The point estimate, $\hat{p}_{.j}$, of $p_{.j}$, $j \in S$ is under the null hypothesis, given by:

$$\hat{p}_{.j} = \frac{n_{.j}}{n_{..}}, \quad (13)$$

where $n_{.j}$ is the sum of transitions to state j for all prior sequences $i \in S^u$ and $n_{..}$ is the total number of transitions to any state for all prior sequences, which is the same as the sample size.

The test statistic for testing the null hypothesis, H_0 , against the alternative hypothesis, H_1 , is given by equation (11) where the v :th order is zero and the u :th order is the optimal order established using BIC. The test statistic under the null hypothesis is asymptotically Chi-square distributed with $(n_s^u - 1)(n_s - 1)$ degrees of freedom.

F. Testing for Time Homogeneity of a Markov Chain

The transition probabilities of the Markov chain is estimated under the assumption of time homogeneity. This assumption has to be validated. A quite intuitive procedure, based on the work of Anderson and Goodman (1957), for testing time homogeneity is to divide the time series into $N > 1$ subintervals of equal length. For time homogeneity to be valid, the TPM must be the same for each of the N subintervals of time. Let subinterval k be denoted by I_k , $k = 1, \dots, N$. Given that the Markov chain of order u has taken the path $i \in S^u$ in subinterval I_k , the transition probability of moving to $j \in S$ is denoted as follows:

$$p_{ij}^k = \mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_{n-u} = i_{n-u}), \quad n \in I_k, i \in S^u, j \in S. \quad (14)$$

The transition probabilities of each subperiod of time are estimated completely analogously to the transition probabilities over the whole time period, using (4) for the subperiods sample. The aim is to test whether or not the TPM for each subperiod is the same

⁷The aim is to test whether or not the probability distribution is the same for the optimal order and the zero-order chains. Let $p_{.j}$ denote the probability of moving to state $j \in S$ regardless of the prior sequence. Mathematically, the null and alternative hypotheses can be described as:

$$\begin{aligned} H_0 &: \forall i \in S^u, \forall j \in S; p_{ij} = p_{.j} \\ H_1 &: \exists i \in S^u, \exists j \in S; p_{ij} \neq p_{.j}. \end{aligned}$$

as the TPM for the whole period. The null and alternative hypotheses can be expressed⁸ as:

$$\begin{aligned} H_0 &: \text{the Markov chain is time homogenous,} \\ H_1 &: \text{the Markov chain is time heterogeneous.} \end{aligned}$$

Under the null hypothesis the likelihood ratio test statistic, Λ , becomes:

$$\Lambda = \prod_{k=1}^N \prod_{i \in S^u, j \in S} \left(\frac{\hat{p}_{ij}^k}{\hat{p}_{ij}} \right)^{n_{ij}^k}. \quad (15)$$

The likelihood ratio test statistic, Λ , is asymptotically equivalent to:

$$-2 \log(\Lambda) = 2 \sum_{k=1}^N \sum_{i \in S^u, j \in S} n_{ij}^k \log \left(\frac{\hat{p}_{ij}^k}{\hat{p}_{ij}} \right). \quad (16)$$

The test statistic $-2 \log(\Lambda)$, under H_0 , is asymptotically Chi-squared distributed with $(N-1)n_s^u(n_s-1)$ degrees of freedom⁹. Here \hat{p}_{ij}^k is the estimate of (14) and \hat{p}_{ij} is the estimate of the transition probability of the u :th order Markov chain over the whole time period, which is given by (4). Since all subintervals are compared to the whole time period the problem of multiple comparisons becomes apparent. Bonferroni's method, by which the significance level is adjusted based on the number of comparisons made, is used.

IV. Data

In this section the data used in this paper is presented and described in detail. Additionally, some summary statistics of the data are given. The data used in this paper is the Nasdaq OMXSPI index, also known as the Stockholm all share index. This index represents the value of all shares that are traded at Stockholm stock exchange (<http://www.nasdaqomxnordic.com>). The price data consists of the closing prices of the index OMXSPI for days, weeks and months respectively (non-trading days are excluded).¹⁰

The index OMXSPI is used as a proxy for the Swedish stock market. The motivation

⁸Let the TPM for the k :th subinterval be denoted by \mathbf{P}^k and the TPM for the whole time period be denoted by \mathbf{P} . The objective is to test whether or not the transition probabilities from each subperiod is the same as the transition probabilities for the whole time period. The null hypothesis and the alternative hypothesis can then mathematically be stated as:

$$\begin{aligned} H_0 &: \forall k \in \{1, \dots, N\}; \mathbf{P}^k = \mathbf{P} \\ H_1 &: \exists k \in \{1, \dots, N\}; \mathbf{P}^k \neq \mathbf{P}. \end{aligned}$$

⁹In appendix C a motivation for this number of degrees of freedom can be found.

¹⁰The data used has been downloaded 2015-04-25 through the Bloomberg terminal.

for using this index over the index OMXS30 is that it includes all traded stocks at the Stockholm Stock Exchange while OMXS30 only consists of the 30 most traded stocks. Therefore, the index OMXSPI serves better as a proxy for the Swedish stock market as a whole than OMXS30 does.

The time period used in this study is January 2000 to April 2015. In particular, for the daily price data, the statistics are based on observations from the time period 2000-01-02 to 2015-04-23. The weekly prices come from the period 2000-01-07 to 2015-04-17 and for the monthly closing prices the period 2000-01-31 to 2015-03-31 has been considered. The returns are calculated as in (1). In table I some descriptive statistics of the samples used throughout this paper are shown¹¹.

Table I
Descriptive Statistics of Prices and Returns.

The table shows some descriptive statistics of daily, weekly and monthly closing prices and the corresponding returns of the index OMXSPI during the period January 2000 - April 2015. In the left part of the table the descriptive statistics of the prices are shown and in the right part the corresponding descriptive statistics of the returns are shown. The statistics shown are the mean, median, standard deviation (Std.), the minimum and maximum value, the inner quantile range (IQR), the skewness, kurtosis and the number of observations (No. obs.).

Descriptive statistics of prices.				Descriptive statistics of returns.			
The descriptive statistics of closing prices during the period January 2000 - April 2015.				The descriptive statistics of returns during the period January 2000 - April 2015.			
	Daily	Weekly	Monthly		Daily	Weekly	Monthly
Mean	302.5067	302.5237	302.6447	Mean	0.0002	0.0011	0.0043
Median	307.3150	306.8850	308.4100	Median	0.0007	0.0040	0.0066
Std.	86.2225	86.3420	86.7245	Std.	0.0144	0.0298	0.0568
Min	126.4100	134.3700	134.3700	Min	-0.0775	-0.2059	-0.1789
Max	560.5500	556.1700	548.6400	Max	0.0901	0.1161	0.1873
IQR	129.3800	128.3900	130.4650	IQR	0.0144	0.0315	0.0568
Skewness	0.1800	0.1721	0.1637	Skewness	0.0793	-0.6785	-0.2641
Kurtosis	2.6567	2.6238	2.6365	Kurtosis	6.5072	7.0059	4.1830
No. obs.	3842	798	183	No. obs.	3841	797	182

V. Test Results

This section presents the results from the various tests we have performed on the Markov chain constructed from the returns of the index OMXSPI. All of these tests are discussed in greater depth in section III, where the Markovian model used in this paper is presented.

¹¹In appendix B plots of time series of prices and returns for the index OMXSPI during January 2000 to April 2015 are given for daily, weekly as well as monthly data.

It is found that the optimal order of the Markov chains representing daily, weekly and monthly returns is 1. This is true both when the benchmark is the geometric return and when it is the zero return. Further, it is found that a random walk behaviour of the Swedish stock market cannot be rejected, nor can the assumption of time homogeneity be rejected, for any of the benchmarks and for all frequencies of returns.

A. The Optimal Order

The highest order allowed for the chain is determined as described in section III.D.1 for each of the frequencies. The highest order allowed is denoted by M in table II. Note that M does not need to be the same for all frequencies of returns. The statistic $-2\log(\Lambda_{u,M})$ (see definition 1) is denoted by $\eta_{u,M}$ to simplify the notation in table II. Further, $f(u)|_M$ denotes the BIC statistic, where order u is tested against order M . In table II these two statistics are displayed for daily, weekly and monthly returns of OMXSPI.

Table II
Test results for the test of the order.

The table shows the test results for the optimal order of the Markov chains, as determined by BIC, describing daily, weekly and monthly returns of the index OMXSPI during the period January 2000 - April 2015. In the first part of the table the benchmark return is the geometric return and in the second part of the table the benchmark is the zero return. The variable $f(u)|_M$ denotes the test statistic calculated using the Bayesian information criterion and the variable $\eta_{u,M}$ is the test statistic calculated in the test of order u against the highest order allowed M .

Geometric return.

The benchmark return used to construct the Markov chain is the geometric return.

	Daily	Weekly	Monthly	Daily	Weekly	Monthly
u	$f(u) _{M=7}$	$f(u) _{M=5}$	$f(u) _{M=3}$	$\eta_{u,M=7}$	$\eta_{u,M=5}$	$\eta_{u,M=3}$
1	-915.3	-175.7	-18.2	124.4	24.5	13.0
2	-902.0	-163.2	-9.7	121.2	23.7	11.0
3	-870.6	-137.4	-	119.6	22.8	-
4	-811.1	-94.4	-	113.0	22.8	-
5	-692.6	-	-	99.5	-	-
6	-470.9	-	-	57.2	-	-

Zero return.

The benchmark return used to construct the Markov chain is the zero return.

	Daily	Weekly	Monthly	Daily	Weekly	Monthly
u	$f(u) _{M=7}$	$f(u) _{M=5}$	$f(u) _{M=3}$	$\eta_{u,M=7}$	$\eta_{u,M=5}$	$\eta_{u,M=3}$
1	-914.5	-175.9	-19.3	125.2	24.3	11.9
2	-900.4	-163.0	-11.2	122.9	24.0	9.6
3	-868.8	-137.6	-	121.4	22.6	-
4	-807.2	-95.7	-	117.0	11.1	-
5	-686.3	-	-	105.9	-	-
6	-470.0	-	-	58.2	-	-

From table II one can see that the function value $f(u)|_M$ is the smallest for $u = 1$ for all the three frequencies of returns for both the benchmarks. Hence, the optimal order of the Markov chain representing daily, weekly and monthly returns is 1. This is true for both benchmarks used to construct the Markov chain modelled.

B. Time Dependence in Returns

With the optimal order established, the test for time dependence is, as described in section III.E, simply a matter of testing a chain of the established optimal order against a chain of

order 0. In this case, the optimal estimate of the order, i.e. the BIC estimate of the order, is 1 for all the three frequencies of returns. This means, for daily, weekly and monthly returns, that the null hypothesis that the Markov chain is of order 0 is tested against the alternative hypothesis that it is of order 1, for each of the frequencies of returns and both benchmarks.

Below in table III, the test results for time dependence in returns are shown. In the left part of the table the test results using the geometric return as the benchmark are shown and in the right part of the table the test results when the zero return is used as the benchmark are shown.

Table III
Test results for time dependence.

The table shows the test results for daily, weekly and monthly data when the optimal order, \hat{u}_{BIC} , established by the Bayesian information criterion, is tested against the order 0. In the left part of the table the benchmark return is the geometric return and in the right part of the table the return is the zero return. Here df is the number of degrees of freedom of the test statistic η in the test for time dependence.

Geometric return.				Zero return.			
The benchmark return used to construct the Markov chain is the geometric return.				The benchmark return used to construct the Markov chain is the zero return.			
	Daily	Weekly	Monthly		Daily	Weekly	Monthly
\hat{u}_{BIC}	1	1	1	\hat{u}_{BIC}	1	1	1
η	0.0701	0.2333	0.3539	η	0.0024	0.0005	1.8321
df	1	1	1	df	1	1	1
p -value	0.7912	0.6291	0.5519	p -value	0.9608	0.9825	0.1759

In both the left and right part of table III, the test results for the Markov chains constructed using the aforementioned benchmarks show high p -values for all three frequencies of returns. At the conventional significance levels (1 %, 5% and 10%) the null hypothesis that the Markov chain of the optimal order is also a Markov chain of order zero cannot be rejected. Hence, we cannot reject that the returns are time independent for any of the three frequencies of returns, and for both benchmarks used to construct the chain.

C. Test for Time Homogeneity

This section presents the test results of the test for time homogeneity outlined in section III.F. Considering the BIC estimate of the order and two subintervals of equal length, time homogeneity cannot be rejected. The details of the test results are presented in table IV below where the benchmark is the geometric return in the left part of the table and the zero return is the benchmark in the right part.

Table IV

Test results for time homogeneity of the chain of the optimal order.

The table shows the test results for time homogeneity of the Markov chains of the optimal order, \hat{u}_{BIC} , representing daily, weekly and monthly returns of the index OMXSPI during the time period January 2000 - April 2015, when the time series of returns is divided into $N = 2$ subintervals of equal length. The test statistic of the test for time homogeneity is denoted by $\eta_{\hat{u}_{\text{BIC}}, N}$. In the left part of the table the benchmark used to construct the chain is the geometric return and in the right table the benchmark is the zero return.

Geometric return.				Zero return.			
The benchmark return used to construct the Markov chain is the geometric return.				The benchmark return used to construct the Markov chain is the zero return.			
	Daily	Weekly	Monthly		Daily	Weekly	Monthly
\hat{u}_{BIC}	1	1	1	\hat{u}_{BIC}	1	1	1
N	2	2	2	N	2	2	2
$\eta_{\hat{u}_{\text{BIC}}, N}$	3.7140	2.6171	1.3805	$\eta_{\hat{u}_{\text{BIC}}, N}$	3.8915	2.0733	1.3455
df	2	2	2	df	2	2	2
p -value	0.1561	0.2702	0.5115	p -value	0.1429	0.3546	0.5103

Here $\eta_{\hat{u}_{\text{BIC}}, N}$ denotes for the test statistic in (16), where \hat{u}_{BIC} is the optimal order established using BIC and N is the number of subintervals of equal length the data set is divided into to perform the test for time homogeneity. For the estimation procedure outlined in section III.C to be valid the chain has to be homogenous. The high p -values show that time homogeneity cannot be rejected for the optimal order for any of the chains representing daily, weekly and monthly returns of OMXSPI if the conventional significance levels are considered. This is true both when the geometric return and the zero return are used as benchmarks.

In table V below, the test results for time homogeneity of the Markov chain of order 0 are shown. In the left part of the table, the test results when the geometric return is used as the benchmark are shown and the test results when the zero return is used as the benchmark are shown in the right part.

Table V

Test results for time homogeneity of the chain of order zero.

The table shows the test results for time homogeneity of the Markov chains of order 0 representing daily, weekly and monthly returns of the index OMXSPI during the period January 2000 - April 2015 when the time series of returns is divided into $N = 2$ subintervals of equal length. The test statistic of the test for time homogeneity is denoted by $\eta_{\hat{u}_{\text{BIC}}, N}$. In the left part of the table the benchmark used to construct the chain is the geometric return and in the right part of the table the benchmark is the zero return.

Geometric return.				Zero return.			
The benchmark return used to construct the Markov chain is the geometric return.				The benchmark return used to construct the Markov chain is the zero return.			
	Daily	Weekly	Monthly		Daily	Weekly	Monthly
u	0	0	0	u	0	0	0
N	2	2	2	N	2	2	2
$\eta_{0,N}$	0.2671	0.3054	0.8571	$\eta_{0,N}$	0.1388	0.2731	3.6785
df	1	1	1	df	1	1	1
p -value	0.6053	0.5805	0.3546	p -value	0.7094	0.6013	0.0551

The p -values corresponding to daily, weekly and monthly returns, in table V are well above any conventional significance level when the geometric return is used as the benchmark. Hence, time homogeneity cannot be rejected for the chains representing the aforementioned frequencies, when the geometric return is used as benchmark. When the zero return is used as the benchmark, time homogeneity cannot be rejected for the Markov chains of order 0 representing daily and weekly returns. The p -value for the Markov chain representing the monthly returns is slightly above 5 %. However, in this case a significance level of 5 % corresponds to an overall significance level of 10 % when Bonferroni's method is used, since two comparisons are made when there are two subintervals. This means that time homogeneity cannot be rejected at the 10 % significance level for any of the three frequencies of returns.

VI. Discussion

The research on market efficiency and random walks in stock prices is as various as it is voluminous. This section aims to place the method developed and the tests performed in this paper in a larger context. It also discusses the reliability of the model used, as well as how the joint hypothesis problem appears in this setting.

The section starts with a brief summary of the results and the conclusions that can be drawn from these, and a validation of the of the assumptions about the Markov chain modelled follows. Thereafter, the joint hypothesis problem is addressed in conjunction

with a discussion on the choosing of benchmarks. The section continues with a comparison with previous research in which random walks in asset prices and weak form efficiency of the Swedish stock market have been evaluated, and thereafter a comparison with studies in which a Markovian approach has been used to test for random walk behaviour in stock prices is made. In particular, the difference in methodology is discussed. A discussion about what this paper brings to the research on market efficiency and random walks in stock prices in general, and to research using a Markovian approach in particular, follows. The section ends with possible extensions of the presented model that have not been implemented within the frame of this paper, but may be of interest for further research.

A. Summary

We find that the random walk hypothesis cannot be rejected for the index OMXSPI for the period January 2000 to April 2015, using the Markovian methodology presented and developed in this paper. This holds true for daily, weekly as well as monthly returns. Further, time homogeneity cannot be rejected for any of the data sets of returns. No evidence that supports that prices can be predicted using historical data is found, which is consistent with that the Stockholm Stock Exchange is weak form efficient.

B. Validation of the Assumptions

The method presented in section III requires the Markov chain representing the stock returns to be aperiodic, irreducible and time homogeneous. If the analysis concerning random walks and efficient markets is to carry any weight, these assumptions need to be validated.

Aperiodicity and irreducibility of the estimated chains are implicitly tested for both in the use of BIC and in the test for time homogeneity, as computations of the test statistics require that the transition probabilities are strictly positive. This is a sufficient condition for aperiodicity and irreducibility of the chain.¹²

Time homogeneity is explicitly tested for, and cannot be rejected for any of the data sets considered in this paper. This increases the reliability of the results as time homogeneity is a necessary condition for the estimation of the TPMs to be valid. However, it should be noted that a failure to reject time homogeneity is not the same as accepting that the chain is time homogeneous; it simply means that, when dividing the chain into a number of subintervals of equal length, the transition probabilities of the subintervals are

¹²It is easy to see that the chain is aperiodic, since whenever all transition probabilities in the TPM are strictly positive each possible path $ij, i \in S^u, j \in S$ can be taken by the chain. As all entries in the TPM are positive, it is possible to move from any previous path $i \in S^u$ to any state $j \in S$, this includes any path that ends with j (i.e. the state which the chain moves to). Therefore it is possible to move from a given state to the same state in one step; hence the period of every state is one and the chain is aperiodic. Also, as all entries are positive it is possible to move to any state from any other state, which means that the chain is irreducible.

not significantly different from the transition probabilities over the whole period. Nevertheless, the failure to reject the null hypothesis does support the assumption of time homogeneity, in the sense that if there would have been a large difference between the TPMs of the different subintervals and the TPM over the entire period, time homogeneity would have been rejected.

C. The Joint Hypothesis Problem and the Choosing of Benchmarks

The joint hypothesis problem, discussed in further detail in section II.A, states that the efficient market hypothesis must be tested jointly with an equilibrium pricing model. This means that to decide whether or not any excess returns can be made, one must first establish a level of returns which can be considered "normal". In this paper, a benchmark return is used to represent the normal return and the simple mapping rule is that any return above the benchmark is classified as "high", and any return below the same benchmark is classified as "low". The benchmarks used to determine in which state to place the return over one time period is the zero return and the geometric return, respectively.

The use of zero as a benchmark is motivated by the fact that any positive return increases the value of a portfolio. Abstracting from reality and considering a risk-free return of zero, investors would prefer to keep their money in the market during such a period. Inversely, a negative return would mean that investors would prefer to stay out of the market. In this setting, it does not matter whether zero is considered to be high or low, as if the return is zero over a period, any investor would be indifferent as to where their money is placed.

A return of zero cannot be said to equal the expected return, as empirical evidence suggests that the expected return of the market should be positive for long time periods. Because the mapping fails to account for magnitudes, it is also impossible to say which periods are the most profitable, or even if a period is more profitable than the average one; all that can be concluded is that the return is positive. Because of this, the benchmark zero is not to be seen as an equilibrium return in an empirical setting, but as a possible equilibrium return in a theoretical abstraction, which can be used to investigate patterns in historical prices.

The other benchmark return that is used is the geometric return, which can hardly be described as the expected return at any time during the time period. Even if the future would be like the past in a probabilistic sense, which would let the geometric return up to a given time point act as a benchmark return for the same point in time, there is no reasonable argument for the opposite. Hence, the geometric return is not to be treated as an equilibrium return, but rather as a benchmark against which the performance of an individual stock can be measured. So, while it cannot be said about returns above the geometric return that they have beaten some kind of expectation, it can be said that they

have performed well relatively to an unbiased average.

D. Comparisons With Other Studies

This section compares our results to previous studies, which are divided into two categories. First, our results are compared to other studies that focused either exclusively or partly on random walk behaviour of the Swedish stock market and Swedish stock market efficiency. Thereafter, a comparison with other studies which have used a Markovian approach is made.

D.1. Comparisons With Other Tests for Swedish Stock Market Efficiency

In section I, four studies that tested the Swedish stock market for random walk behaviour in stock prices and weak form efficiency are discussed. Three of these studies concerned random walks in stock prices (Jennergren & Korsvold, 1974; Frennberg & Hansson, 1993; Shaker, 2013). The exception is the study by Metghalchi, Chang and Marcucci (2008), which tested for the profitability of three trading rules based on moving averages.

Metghalchi, Chang and Marcucci (2008) found that trading rules based on moving averages can be profitable even when transaction costs are accounted for, which violates both weak form market efficiency and that stock prices follow a random walk. These results stand in contrast to the results obtained using the Markovian approach in our paper, which suggest that a random walk behaviour of the Swedish stock market cannot be rejected. One possible explanation for the difference in results is that the time period used differs. As mentioned in section I Metghalchi, Chang and Marcucci used data from the period 1986 to 2004. As of today, when the computer technology is developed to a large extent, such profit opportunities are more likely to vanish rapidly, as high-frequency robots exploit such opportunities instantly.

In contrast to our results studies by Jennergren and Korsvold (1974) as well as Frennberg and Hansson (1993) reject random walks in stock prices of stocks traded at the Swedish stock market. The difference in results can be explained by several factors. The time series used is not the same since our sample is from the 21st century while their samples are from the 20th century. It is possible that the market was less random prior to the development of fast computer communications. Furthermore, the methodology used in this paper differs from the ones used in these two papers. In addition, Jennergren and Korsvold did use individual stocks traded at the Swedish stock exchange while we use an index as a proxy for the Swedish stock market.

A comparison to Shaker (2013) is especially appealing since Shaker uses the index OMXS30 from January 2003 to January 2013 which is very similar to OMXSPI for the period January 2000 to April 2015, which is used in this paper. Shaker rejected random walk behaviour for this time period while we conclude that a random walk in the price of the index OMXSPI cannot be rejected. The Markovian model that is used in the context of

this paper differs from the linear model Shaker used to test for serial correlation in returns. Since these two indices are very similar and the time period Shaker used is included in our sample's time period the difference in results is surprising. One possible explanation could be that some assumptions in either our Markovian model or in Shakers linear model have been violated. However, none of the assumptions made in the Markovian model in this paper can be rejected. Another possible explanation can be the choosing of benchmark in the construction of the Markov chain. Other benchmarks, which are not considered in this paper, might give different results.

D.2. Comparisons With other Markovian Studies

There are other papers that have used Markovian models for testing random walk behaviour in stock prices, e.g. those presented in section I (cf. Niederhoffer & Osborne, 1966; Fielitz & Bhargava, 1973; Fielitz, 1975; McQueen & Thorley, 1991; Tan & Yilmaz, 2002). This section goes through some of the important differences of our paper as compared to earlier papers that have used Markovian models.

The main difference between the Markovian model in this paper and the Markovian models in earlier papers is how the order is established. McQueen and Thorley (1991), Fielitz and Bhargava (1973), and Fielitz (1975) chose a model with a specified order, they did not test if this order was optimal or not. Nevertheless, it should be mentioned that McQueen and Thorley presented some arguments supporting their choice to use a second-order markovian model. Our model on the other hand, does not make any a priori assumption regarding the order of the chain, instead an optimal order is derived. Tan and Yilmaz (2002) did address how to determine which order to use. However, the procedure they suggested for determining the order is incorrect. Since every Markov chain of order u is a chain of order $u + 1$ as well, it is not possible to make the pairwise tests of orders as they suggested. In our paper BIC is used to get around this problem. Both the consistency and the optimality of the BIC estimator give additional support for the usage of BIC.

Further, some papers that have used a Markovian model for testing random walk behaviour in stock prices have made crucial assumptions that were never tested for. McQueen and Thorley (1991) assumed time homogeneity of the Markov chain representing the returns of the NYSE, without testing for it. Fielitz and Bhargava (1973) on the other hand were aware of the importance of time homogeneity. They tested for it, rejected it, and still proceeded with the analysis and rejected random walk behaviour of the stocks considered. The assumption of time homogeneity is tested and cannot be rejected for the samples considered in this paper, which is the best possible outcome since the null hypothesis is must be that the chain is time homogenous.

Tan and Yilmaz (2002) criticised McQueen and Thorley (1991) for not testing the assumption of time homogeneity. According to Tan and Yilmaz the assumption of time homogeneity would have been rejected for the data set McQueen and Thorley (1991)

used. Nonetheless, the null distribution of the test statistic Tan and Yilmaz suggested had the wrong degrees of freedom (see appendix C for the correct degrees of freedom). This paper uses an approach first suggested by Anderson and Goodman (1957) for testing the assumption of time homogeneity, which is similar to the approach used by Fielitz and Bhargava (1973), and Fielitz (1975). However, our results can be reproduced as the number of subintervals used is stated explicitly, which is not the case in the papers by Fielitz and Bhargava, and Fielitz.

E. Contributions

This paper offers an alternative approach for testing for a random walks in stock prices using a Markovian model to capture dependence structures in returns, which may be non-linear. One big advantage of this model compared to others, is that the model is nonparametric, i.e. no assumptions about the distribution of returns have to be made. The main contribution to the field of market efficiency and random walk theory is the development of an existing method for testing for random walk behaviour of a stock market, in the sense that this paper provides a way to optimally determine the order of the Markov chain modelled. Furthermore, it does contribute with a crucial correction of the asymptotic distribution of the test statistic used in the test for time homogeneity in earlier studies. From an economic point of view, this paper contributes with test results on random walk behaviour in prices, using a method which, as far we know, never has been applied to the Swedish stock market.

F. Suggestions for Further Studies

One of the restrictions within the scope of this paper is that the state space consists of only two states. Future studies may use the method presented in this paper, but include more states, which would allow the model to capture not only the directions, but also, to some degree, the magnitudes of the returns. Another suggestion would be to use two states, one state that represents positive returns and one that represents negative returns, and in addition to these states include a variable that predicts the magnitude of a return, given that it is positive, or negative, and has followed a particular path. Such a model has been used by Lennartsson, Baxevani and Chen (2008) to capture the amount of precipitation in Sweden.

The model can be extended to a vector process (cf. Fielitz & Bhargava, 1973), which considers all firms traded at the Stockholm Stock Exchange, or another market, simultaneously. This extension can use BIC for order selection, allowing for different orders among the firms used in the vector process. A combination of this procedure and a chain that allows for more than two states is also a possibility.

Appendix A. Markov Theory

Basic theory on Markov chains is presented in this section. It starts with the definition and some properties of first-order Markov chains (in the first subsection simply referred to as Markov chains), and then extends the definition and properties of first-order Markov chains to higher-order Markov chains.

First-Order Markov Chains

Consider a set of states $S = \{s_1, s_2, \dots\}$, henceforth referred to as a state space, and a discrete time random process $\{X_n : n \in \mathbb{N}\}$ that moves, or transitions, from one state in the state space to another. The process is called a Markov chain if the probability distribution of the future state is independent of all previous states except for the current one. The formal mathematical definition of a Markov chain is given below:

DEFINITION 2: *Let S be a countable state space. The process $\{X_n : n \in \mathbb{N}\}$ is a Markov chain if it satisfies the Markov property:*

$$\mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_n = j | X_{n-1} = i_{n-1}) \quad (\text{A1})$$

$$\forall n \geq 1, \forall i_0, \dots, i_{n-1}, j \in S.^{13}$$

This definition, as well as others presented in this subsection, is based on the notations and terminology used by Grimmet and Stirzaker (2001).

The transition probabilities and the transition probability matrix (TPM) of the Markov chain $\{X_n : n \in \mathbb{N}\}$, henceforth denoted by X , is defined as:

DEFINITION 3: *Let S be a countable state space and X a discrete time Markov chain. The transition probability from state i in step $n - 1$ to state j in step n is denoted $p_{ij}(n - 1, n) = \mathbb{P}(X_n = j | X_{n-1} = i)$. The transition probability matrix $\mathbf{P}(n - 1, n) = (p_{ij}(n - 1, n))$ is the $n_s \times n_s$ matrix of transition probabilities $p_{ij}(n - 1, n)$, where n_s denotes the cardinality of the state space.¹⁴*

For the purpose of further reference, an important property of Markov chains is irreducibility, which mathematically is defined as:

DEFINITION 4: *Let X be a Markov chain defined on state space S . The chain X is said to be irreducible if:*

$$\forall i, j \in S, \exists m \in \mathbb{Z}_+, m < \infty : \mathbb{P}(X_{n+m} = j | X_n = i) > 0. \quad (\text{A2})$$

¹³Every i_{n-k} , $k = 1, \dots, n$ equals some state $s_l \in S$, $l = 1, 2, \dots$

¹⁴The cardinality of a state space S is commonly denoted by $|S|$. To simplify notation, especially in the method section, n_s will be used throughout this paper.

Another important property of Markov chains is aperiodicity, which is related to the period of the chain. Both concepts are defined below:

DEFINITION 5: *Let X be a Markov chain defined on state space S . The state i is said to have period d_i , where d_i is:*

$$d_i = \gcd\{m : \mathbb{P}(X_m = i | X_0 = i) > 0\}, \quad (\text{A3})$$

where \gcd stands for greatest common divisor. A state is said to be aperiodic if $d_i = 1$.

If the probability of a transition from state i to j does not depend on when the chain is in state i or j the chain X is called time homogenous. Formally this can be defined as:

DEFINITION 6: *The Markov chain X over the state space S is called time homogenous if*

$$p_{ij}(n-1, n) = p_{ij}(0, 1) \quad (\text{A4})$$

$$\forall n \geq 1, \forall i, j \in S.$$

For a time homogenous chain the notation p_{ij} is used to denote the probability for each one-step transition from i to j , thus $p_{ij}(0, 1) = p_{ij}$.

Let X be a time homogenous Markov chain defined on a state space S with L states. The transition probability matrix (TPM), here denoted by \mathbf{P} ,¹⁵ can then be stated as follows:

$$\mathbf{P} = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1L} \\ p_{21} & p_{22} & \cdots & p_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ p_{L1} & p_{L2} & \cdots & p_{LL} \end{pmatrix}. \quad (\text{A5})$$

Higher-Order Markov Chains

Higher-order Markov chains can be seen as generalisations of first-order Markov chains. The order refers to the number states prior to the future one that may carry information about the future outcome. The definitions given below are straight forward generalisations from the ones concerning first-order Markov chains. Formally, a Markov chain of order u is defined as follows:

DEFINITION 7: *Let $S = \{s_1, s_2, \dots\}$ be an at most countable state space and $\{X_n, n \in \mathbb{N}\}$*

¹⁵In the matrix given in (A5), 1 represents the state $s_1 \in S$ and 2 represents $s_2 \in S$. Analogously each positive integer k represents $s_k \in S$. Note that the state space is finite with cardinality L .

be a discrete-time stochastic process. Then $\{X_n, n \in \mathbb{N}\}$ is a Markov chain of order u if:

$$\mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_{n-u} = i_{n-u}) \quad (\text{A6})$$

$$\forall n \geq u, \forall j, i_{n-1}, \dots, i_{n-u}, \dots, i_0 \in S.$$

By this definition, a first-order Markov chain is also a second-order Markov chain. In fact it follows directly from the definition that a Markov chain of order u is also a Markov chain of order $u + 1$.¹⁶ In other words it is a sufficient, but not necessary, condition for a Markov chain of order $u + 1$ to be a Markov chain of order u .

REMARK 1: *It is consistent with the discussion above to think about a Markov chain of order zero. As an example consider any sequence of independent random variables, that takes values in a countable set.*¹⁷

The definitions of transition probabilities and the transition probability matrix (TPM) as well as concepts such as time homogeneity are defined analogously to those for a Markov chain of order one. For reference purposes these definitions can be found below.

DEFINITION 8: *Let $S = \{s_1, s_2, \dots\}$ be an at most countable state space and $S^u = \{s_{n_1} \dots s_{n_u} : \forall s_{n_k} \in S\}$ be the state space containing all possible sequences of length u consisting of states $s_n \in S$. Consider a u :th order Markov chain X . The transition probability $p_{ij}(n - u, n)$ to end up in $j \in S$ at time n after having followed the path described by the sequence $i \in S^u$ is defined as:*

$$p_{ij}(n - u, n) = \mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_{n-u} = i_{n-u}) \quad (\text{A7})$$

$$i = i_{n-u} \dots i_{n-1} \in S^u, \quad j \in S.$$

The transition probability matrix $\mathbf{P}(n - u, n) = (p_{ij}(n - u, n))$ is then the $n_s^u \times n_s$ matrix of transition probabilities $p_{ij}(n - u, n)$.

Note that as a probability is assigned to each combination of previous states, the transition probability matrix is no longer a square matrix, unless $u = 1$. As stated in definition 8, the states in the chain prior to the future one belongs to the state space S^u which consists of all possible sequences, of length u , of states in S . This means that for a second-order chain with only two states, s_1 and s_2 , the state space of interest is $S^2 = \{s_1 s_1, s_1 s_2, s_2 s_1, s_2 s_2\}$. Note that $s_1 s_2$ and $s_2 s_1$ represents different sequences; the

¹⁶See appendix C for a motivation.

¹⁷Assume that X_1, X_2, \dots are independent variables taking values in some countable set. Then:

$$\mathbb{P}(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_1 = x_1) = \mathbb{P}(X_n = x_n),$$

where the equality follows from the independence of the random variables. This is a Markov chain of order zero.

first one represents that the chain moves from s_1 to s_2 and the second one represents the reversed movement.

For a Markov chain of order u irreducibility and aperiodicity are defined analogously to the first-order chains. A chain is irreducible if all states are accessible from each other, i.e the probability of moving from one state $i \in S$ to another state $j \in S$ is positive in any finite number of transitions. It follows that irreducibility is independent of the order. Furthermore, the period, d_i , of an u :th order chain is the greatest common divisor of the possible paths that can be taken from one state $s_k \in S$ to the same state $s_k \in S$. If $d_i = 1$ then the u :th order chain is aperiodic. In particular, if $p_{ij} > 0$ for all $i \in S^u, j \in S$, then the chain is aperiodic.

The Markov chain is time homogenous if the transitions following a certain path depend only on the sequence of states, and not on when the sequence starts. Formally, this is defined:

DEFINITION 9: *The Markov chain X defined on the state space S is called time homogenous if*

$$p_{ij}(n - u, n) = p_{ij}(0, u) \quad (\text{A8})$$

$$\forall n \geq u, \forall i \in S^u, \forall j \in S.$$

For a time homogenous chain the notation p_{ij} is used to denote the probability for each transition following the sequence i to j .

REMARK 2: *If the sequence considered in remark 1 is identically distributed the Markov chain is time homogenous* ¹⁸.

Appendix B. Figures

In the figures below, the prices of the index OMXSPI and the corresponding returns are displayed for monthly, weekly and daily data, respectively. The purpose of these figures is to give a general idea of whether or not the prices follow a random walk.

¹⁸If the sequence X_1, X_2, \dots of random variables considered in remark 1 are identically distributed in addition to independently distributed. Then: $\mathbb{P}(X_n = x_n)$ is the same for all n since the probability distribution is identical for all random variables.

Figure 2.
Plots of monthly prices and returns.

The figure shows plots of the monthly closing prices and the returns of the index OMXSPI during the period 2000-01-01 to 2015-03-31. In the first plot, the price is displayed at the vertical axis and the number of the month is displayed at the horizontal axis. In the second plot, the return is displayed at the vertical axis and the number of the month is displayed at the horizontal axis.

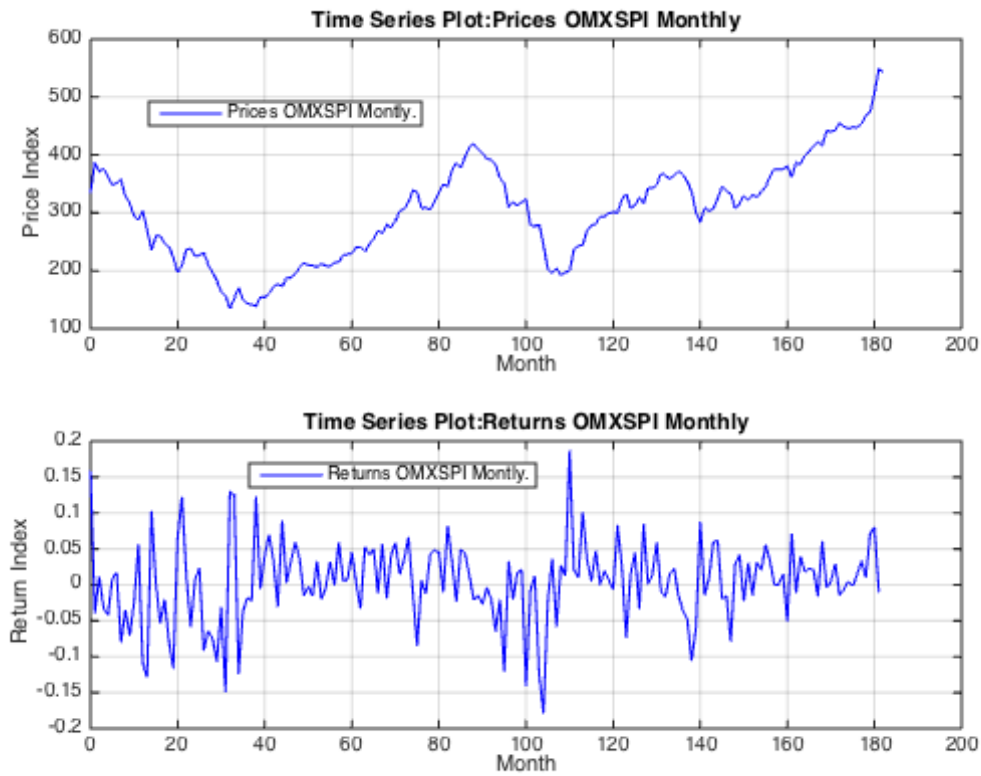


Figure 3.
Plots of weekly prices and returns.

The figure shows plots of the weekly closing prices and the returns of the index OMXSPI during the period 2000-01-01 to 2015-04-17. In the first plot, the price is displayed at the vertical axis and the number of the week is displayed at the horizontal axis. In the second plot, the return is displayed at the vertical axis and the number of the week is displayed at the horizontal axis.

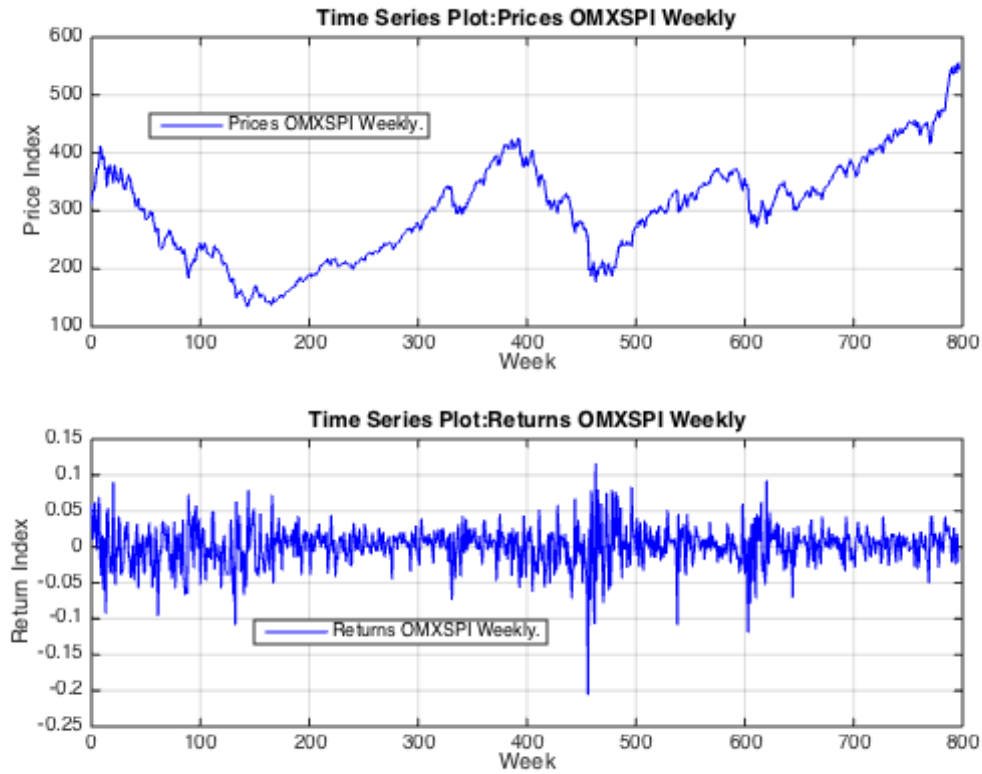
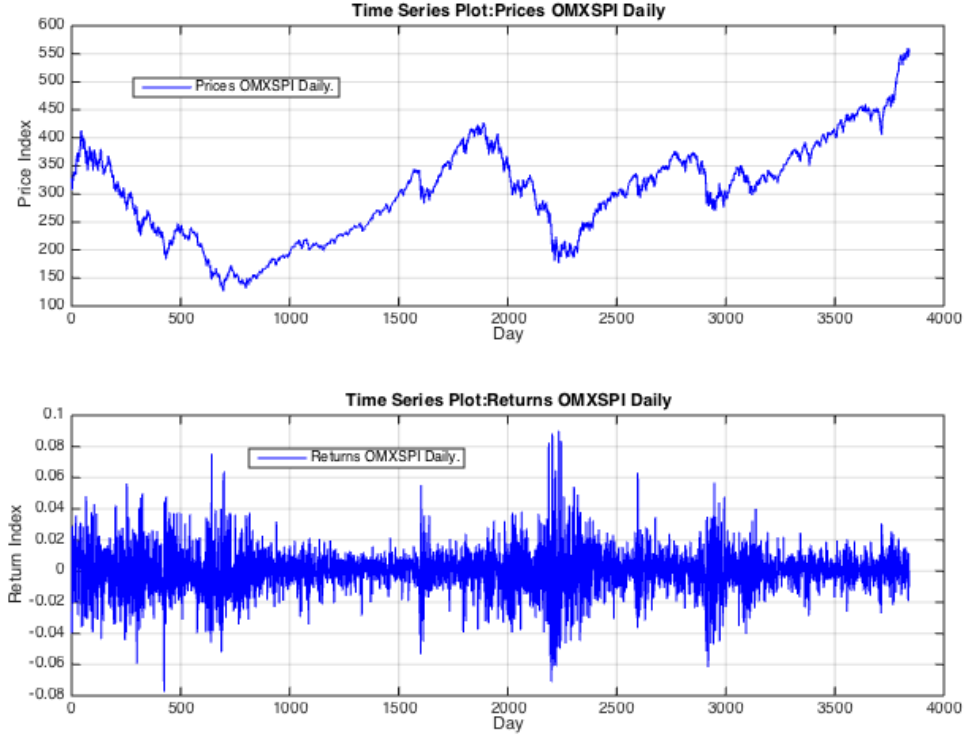


Figure 4.
Plots of daily prices and returns.

The figure shows plots of the daily closing prices and the returns of the index OMXSPI during the period 2000-01-01 to 2015-04-23. In the first plot, the price is displayed at the vertical axis and the number of the day is displayed at the horizontal axis. In the second plot, the return is displayed at the vertical axis and the number of the day is displayed at the horizontal axis.



Appendix C. Miscellaneous

C1 - Derivation of the MLEs of a Transition Probability

Let X be a time homogenous Markov chain of order u on a state space S . Define S^u as the state space consisting of all possible sequences in u steps on S . Let Y_1, \dots, Y_T be independent random variables such that Y_l takes any value corresponding to the possible sequences, $ij, i \in S^u, j \in S$, that is the observed value of $Y_l, y_l = ij$. Then the probability of $Y_l = y_l$ is:

$$\mathbb{P}(Y_l = y_l) = p_{ij} \quad , i \in S^u, j \in S,$$

T is the sample size. The likelihood function, L , can then be written as:

$$L = \mathbb{P}(Y_1 = y_1, \dots, Y_T = y_T) = \prod_{l=1}^T \mathbb{P}(Y_l = y_l) = \prod_{i \in S^u, j \in S} p_{ij}^{n_{ij}}, \quad (\text{C1})$$

where p_{ij} is the transition probability of a u :th order time homogenous chain from state $i \in S^u$ to $j \in S$, n_{ij} is the number of transitions from $i \in S^u$ to $j \in S$ observed in the times series used. Furthermore, define n_i as the total number of times the chain was observed in state $i \in S^u$. The log-likelihood function, l , is defined as the natural logarithm, \log , of L :

$$l = l(p_{ij}, i \in S^u, j \in S) = \sum_{i \in S^u, j \in S} n_{ij} \log(p_{ij}). \quad (\text{C2})$$

The objective is to maximise (C2) subject to the constraints:

$$\sum_{j \in S} p_{ij} = 1, \quad p_{ij} \geq 0, \forall i \in S^u, j \in S. \quad (\text{C3})$$

Let \mathcal{L} be the Lagrangian function, then the objective is to find the maximum of the Lagrangian:

$$\mathcal{L} = \sum_{i \in S^u, j \in S} n_{ij} \log(p_{ij}) + \lambda \left(1 - \sum_{j \in S} p_{ij} \right). \quad (\text{C4})$$

If the cardinality is n_s , then there are n_s^u states in S^u . Hence, when maximising the Lagrangian, there are $n_s^u \cdot n_s + 1$ first order conditions. For all $i \in S^u$ and all $j \in S$ the following condition holds:

$$\frac{\partial \mathcal{L}}{\partial p_{ij}} = 0 \implies p_{ij} = \frac{n_{ij}}{\lambda}, \forall i \in S^u, j \in S. \quad (\text{C5})$$

By taking the partial derivate w.r.t. the Lagrangian multiplier, λ , the first order condition becomes:

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 \implies 1 = \sum_{j \in S} p_{ij}, \quad (\text{C6})$$

by using (C5) in (C6) the equation can be solved for λ :

$$1 = \sum_{j \in S} p_{ij} = \sum_{j \in S} \frac{n_{ij}}{\lambda} = \frac{n_{i.}}{\lambda} \iff \lambda = n_{i.}, \quad (\text{C7})$$

by substituting (C7) back to (C5) the maximum likelihood estimate, \hat{p}_{ij} of p_{ij} becomes:

$$\hat{p}_{ij} = \frac{n_{ij}}{n_{i.}}, \quad \forall i \in S^u, j \in S, \quad (\text{C8})$$

which is the maximum likelihood estimate of the u :th order Markov chain transition probability, as given in equation (4) in section III.C.

C2 - A Markov Chain of Order u is a Chain of Every Higher Order

Assume that X is a Markov chain of order u . By the definition of a Markov chain of order u :

$$\mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_{n-u} = i_{n-u}).$$

Since the random variable $X_{n-(u+1)}$ does not carry any information according to the Markov property, it follows that:

$$\mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_{n-(u+1)} = i_{n-(u+1)}) = \mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_{n-u} = i_{n-u}).$$

The equalities above yield the following result:

$$\mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_n = j | X_{n-1} = i_{n-1}, \dots, X_{n-(u+1)} = i_{n-(u+1)}).$$

Which means that X is a Markov chain of order $u + 1$. By induction it follows that X is a Markov chain of order $u + v$ for all $v \in \mathbb{N}$: $u + v \leq n$, where $n \in \mathbb{N}$.

C3 - Generating a Markov Chain of a Given Order

Consider a Markov chain of order u with the state space $S = \{s_1, s_2\}$, the starting distribution $\pi = [\frac{1}{2} \ \frac{1}{2}]$ and the TPM

$$\mathbf{P} = \begin{pmatrix} p_{s_1 k_1 s_1} & p_{s_1 k_1 s_2} \\ p_{s_1 k_2 s_1} & p_{s_1 k_2 s_2} \\ \vdots & \vdots \\ p_{s_1 k_n s_1} & p_{s_1 k_n s_2} \\ p_{s_2 k_1 s_1} & p_{s_2 k_1 s_2} \\ \vdots & \vdots \\ p_{s_2 k_n s_1} & p_{s_2 k_n s_2} \end{pmatrix}, \quad (\text{C9})$$

where $k_j \in S^{u-1}$, $j = 1, \dots, n$. Let $p_{s_1 k_j s_1} = p$ and $p_{s_2 k_j s_1} = 1 - p$ for all $j = 1, \dots, n$, resulting in the TPM

$$\mathbf{P} = \begin{pmatrix} p & 1 - p \\ p & 1 - p \\ \vdots & \vdots \\ p & 1 - p \\ 1 - p & p \\ \vdots & \vdots \\ 1 - p & p \end{pmatrix}, \quad (\text{C10})$$

where each entry in the matrix above corresponds to the same entry in the previous one. Note that this construction of the TPM results in a chain whose transition probabilities depend only on the realisation of the stochastic variable observed u periods earlier. To emphasise, the path of the chain does depend on the state the chain was in u periods earlier but, by assumption, not on any other realisations. Hence the process is a Markov chain of order u .¹⁹

Now consider the same process and the construction of the TPM if only $u - 1$ periods prior to the current one are considered. The TPM can be written as:

$$\mathbf{P} = \begin{pmatrix} p_{k_1 s_1} & p_{k_1 s_2} \\ p_{k_2 s_1} & p_{k_2 s_2} \\ \vdots & \vdots \\ p_{k_n s_1} & p_{s_1 k_n s_2} \end{pmatrix}, \quad (\text{C11})$$

where k_j , $j = 1, \dots, n$ is the same as above. Because of the symmetry of the starting distribution as well as the the TPM of the chain of order u , each entry in the matrix

¹⁹Trivially, the process is also a Markov Chain of any order w such that $w > u$ (see appendix C for further details).

above equals the average of the transition probabilities corresponding to the sequences $p_{s_1 k_j s_i} = p$ and $p_{s_2 k_j s_i} = 1 - p$, $j = 1, \dots, n$, $i = 1, 2$, which equals $\frac{p+1-p}{2} = \frac{1}{2}$, resulting in the TPM

$$\mathbf{P} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \\ \vdots & \vdots \\ 1/2 & 1/2 \end{pmatrix}, \quad (\text{C12})$$

in which the probability of a transition to s_1 and s_2 respectively is obviously $\frac{1}{2}$ regardless of which path the chain has taken previously. In other words, the chain is independent of every realisation up to $u - 1$ periods back. As the chain actually does depend on the realisation u periods back, it is not a Markov chain of order $u - 1$. It is, however, impossible to determine this from the TPM above. Indeed, every entry of every TPM where $v, v < u$, previous periods are considered will equal $\frac{1}{2}$. This means that if the chain is (incorrectly) assumed to be of order $u - 1$, the TPM when $u - 1$ previous periods are considered would tell us that the process is a random walk. This is obviously not the case, as the chain is, by construction, a Markov chain of order u .

C4 - Degrees of Freedom in Test for Time Homogeneity

The chi-square distributed test statistic used to determine the reasonableness of the time homogeneity assumption is calculated by summing the logarithmic differences between the estimated transition probabilities for each subperiod and the corresponding estimated transition probabilities for the entire period. The null hypothesis that the Markov chain is time homogenous is then rejected if the probability of finding a test statistic as extreme as the calculated one, given that the null hypothesis is true, is sufficiently small; i.e. we reject the null if the observed p -value is lower than the chosen significance level α . In this setting, the previous question translates to whether or not the estimated transition probabilities of the subintervals differ enough from the transition probabilities over the entire time period to reject the null hypothesis that the transition probabilities are the same over the whole time period. In order to determine whether this is the case or not one must know the distribution of the test statistic. According to Anderson and Goodman (1957) the statistic of interest is chi-squared distributed. To be more precise, it follows a distribution which belongs to the family of chi-squared distributions. Which chi-squared distribution it follows is determined by the degrees of freedom of the test statistic.

The degrees of freedom, df , of a test statistic is defined as the number of values, observed or estimated, which are used in the calculation of the test statistic and may vary freely. Let us consider what this means for the test statistic in the test for time homogeneity. The data is divided into N subintervals, and for each of these a TPM

is estimated. The TPM over the entire period is then simply the weighted average of the TPMs for each subinterval, which means that once the transition probabilities for the $N - 1$ first subintervals have been estimated, the TPM of the N :th subinterval must be such that the weighted average of all of them is the TPM for the entire period. In other words, the TPMs of $N - 1$ subintervals may vary freely. Each TPM consists of n_s^u rows, where n_s^u is the cardinality of the state space S^u ; and n_s columns, where n_s is the cardinality of the state space S . Each row can be seen as representation of a multinomial distribution with n_s outcomes. As the probabilities of each row must sum to one, all but one of the transition probabilities may vary freely. With $N - 1$ subintervals in which the TPMs may vary freely, and $n_s^u(n_s - 1)$ transition probabilities that may vary freely in each TPM, the definition of degrees of freedom, df , gives:

$$df = (N - 1)n_s^u(n_s - 1), \quad (\text{C13})$$

degrees of freedom for the test statistic (16) of time homogeneity.

REFERENCES

- Anderson, T. W., & Goodman, L. A. (1957). Statistical Inference About Markov Chains. *The Annals of Mathematical Statistics*, 28(1), 89-110.
- Fama, E. F. (1965). Random Walks in Stock Market Prices. *Financial Analysts Journal*, 21(5), 55-59.
- Fama, E. F. (1970). Efficient Capital Markets: a Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383- 417.
- Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance*, 46(5), 1575-1617.
- Fama, E. F. (2014). Two Pillars of Asset Pricing. *American Economic Review*, 104(6), 1467-1485.
- Fama, E. F., & Blume, M. E. (1966). Filter Rules and Stock Market trading. *Journal of Business*, 39(1), 226-241.
- Fama, E. F., & French, K. R. (1988). Permanent and Temporary Components of Stock Prices. *Journal of Political Economy*, 96(2), 24-73.
- Fielitz, B. D. (1975). On the Stationarity of Transition Matrices of Common Stocks. *Journal of Financial and Quantitative Analysis*, 10(2), 327-339.
- Fielitz, B. D., & Bhargava, T. N. (1973). The Behavior of Stock-Price Relatives: A Markovian Analysis. *Operations Research*, 21(6), 1183-1199.
- Frennberg, P., & Hansson, B. (1993). Testing the random walk hypothesis on Swedish stock prices: 1919-1990. *Journal of Banking & Finance*, 17(1), 175-191.
- Grimmett, G., & Stirzaker, D. R. (2001). *Probability and random processes*. (3. ed.) Oxford: Oxford University Press.
- Jennergren, P. L., & Korsvold, P. E. (1974). Price Formation in the Norwegian and Swedish Stock Markets: Some Random Walk Tests. *The Swedish Journal of Economics*, 76(2), 171-185.
- Jensen, M. J. (1978). Some anomalous evidence regarding market efficiency. *Journal of Financial Economics*, 6(2), 95-101.
- Katz, R. W. (1981). On Some Criteria for Estimating the Order of a Markov Chain. *Technometrics*, 23(3), 243-249.
- Lennartsson, J., Baxevani, A., & Chen, D. (2008). Modelling precipitation in Sweden using multiple step markov chains and a composite model. *Journal of Hydrology*, 363(1-4), 42-59.
- Lo, A. W., & MacKinlay, A. C. (1988). Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test. *Review of Financial Studies*, 1(1), 41-66.
- Lo, A. W., & MacKinlay, A. C. (2002[1999]). *A non-random walk down Wall Street*. Princeton, N.J.: Princeton University Press.
- Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *The Journal of Economic Perspective*, 17(1), 59-82.

- Malkiel, B. G. (2005). Reflections on the efficient market hypothesis: 30 years later. *The financial review*, 40(1), 1-9.
- Metghalchi, M., Chang, Y-H., & Marcucci, J. (2008). Is the Swedish stock market efficient? Evidence from some simple trading rules. *International Review of Financial Analysis*, 17(3), 475-490.
- McQueen, G., & Thorley, S. (1991). Are Stock Returns Predictable?: A Test Using Markov Chains. *The Journal of Finance*, 46(1), 239-263.
- Niederhoffer, V., & Osborne, M. F. M. (1966). Market making and Reversal on the Stock Exchange. *Journal of the American Statistical Association*, 61(316), 897-916.
- Poterba, J. & Summers, L. (1988). Mean Reversion in Stock Prices: Evidence and Implications. *Journal of Financial Economics*, 22(1), 27-59.
- Schiller, R. J. (2014). Speculative Asset Prices. *American Economic Review*, 104(6), 1486-1517.
- Schwartz, G. E. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, 6(2), 461-464.
- Shaker, A. T. M. (2013). Testing the Weak-Form Efficiency of the Finnish and Swedish Stock Markets. *European Journal of Business and Social Sciences*, 2(9), 176-185.
- Tan, B., & Yilmaz, K. (2002). Markov Chain Test for Time Dependence and Homogeneity: An Analytical and Empirical Evaluation. *European Journal of Operational Research*, 137(3), 524-543.

